

This application is a continuation of U.S. Patent Application Serial No. 09/768,947, entitled "DETECTING DUPLICATE AND NEAR-DUPLICATE FILED", filed on January 24, 2001 and listing William Pugh and Monika H. Henzinger as inventors. Benefit to that application is claimed under 35 U.S.C. §§ 120. That application is expressly incorporated herein by reference. --

IN THE CLAIMS:

Please amend the claims as follows:

Please cancel claims 1-45 without prejudice to, or disclaimer of, the subject matter recited therein.

Please add the following new claims:

- 1 46. (NEW) A method for filtering search results to remove  
2 near-duplicates, the method comprising:
  - 3 a) for each of a predetermined number of candidate  
4 search results, determining whether the candidate  
5 search result is a near-duplicate of another candidate  
6 search result by
    - 7 1) comparing a cluster identifier of the  
8 candidate search result with that of the other  
9 candidate search result, and
    - 10 2) if the cluster identifiers of the two  
11 candidate search results match, then concluding  
12 that the two candidate search results are  
13 near-duplicates; and